



The Mask Dispensers Project



Project goals:

- Generally, coronavirus propagates more in the places where there is a big density of people
- Masks provides the protection needed to stop and control the propagation
- However, some either lose their masks or overused them
- The goal is to cluster paris neighborhoods in such a way we can plant mask dispensers in an equal way taking account of the more critical places.



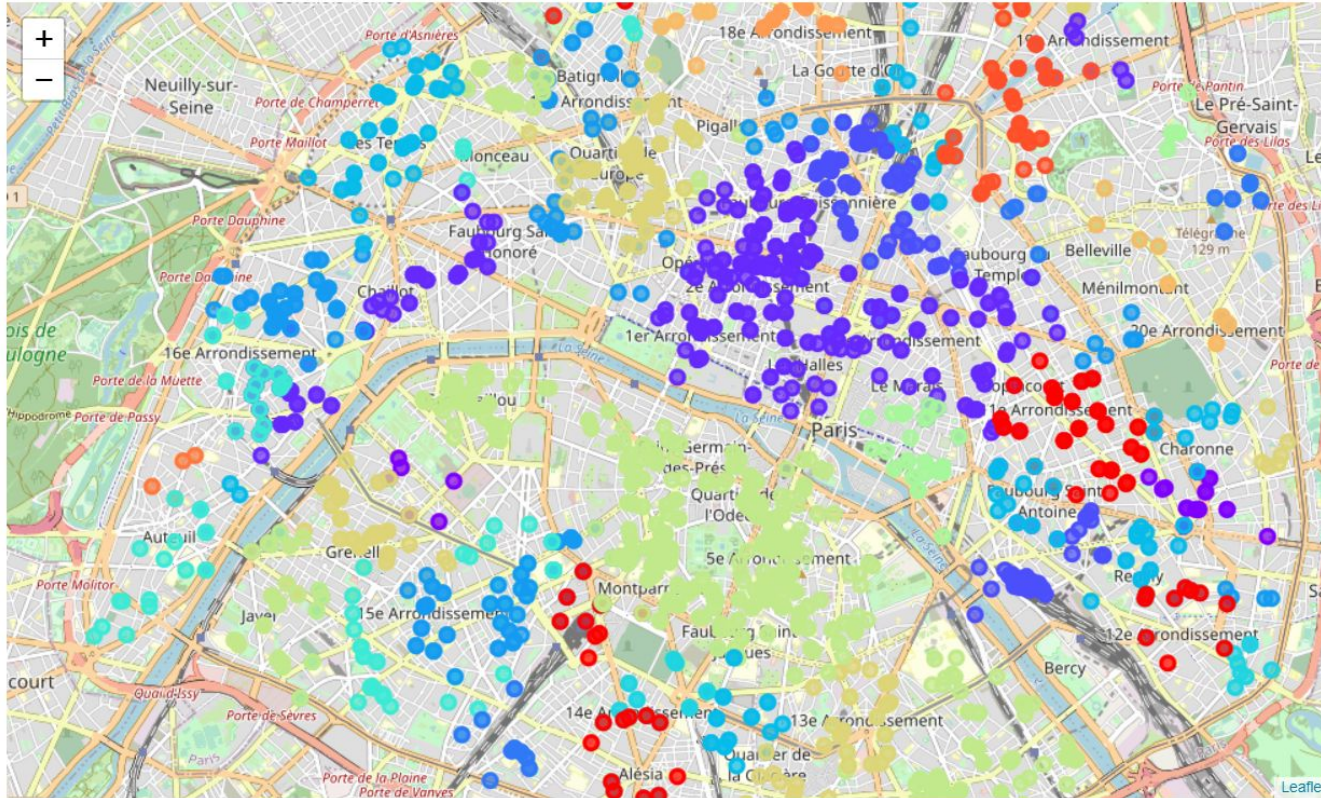
Data acquisition and cleaning:

- Our first source of data is a csv file which contains the medical labs located in Paris.
- This csv already contains the coordinates however we needed to process that column for a better understand
- The second source of data is a Foursquare API, which we precise the category as "school, universities, train stations and hospitals".
- And last, is a web page which we scrapped to get the 20 arrondissement of Paris that we added coordinates to it using GeoCoder API.

Out[82]:

	Cluster Labels	Neighbourhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	3	10ème - Boulevard de Magenta	Platform	Trade School	High School	Train Station	Middle School	College & University	College Arts Building	University	Medical School	Nursery School
1	3	10ème - Rue La Fayette	Platform	Train Station	Train	High School	University	Trade School	College & University	Middle School	College Classroom	Hospital
2	3	10ème - Rue de Dunkerque - Gare du Nord	Platform	Train	Train Station	High School	Hospital	Trade School	Middle School	University	College & University	College Classroom
3	6	10ème - Rue du Faubourg Saint Martin	Trade School	Platform	College Academic Building	Train Station	Hospital	University	General College & University	Elementary School	College Classroom	College Cafeteria
4	0	11ème - Avenue Ledru Rollin	Elementary School	College Arts Building	Music School	College Library	Cooking School	College Academic Building	Middle School	Private School	College Stadium	College Math Building

- Schools, colleges and trains are the more critical venue.



Some of the results obtained were good, however we can notice that some points in some clusters are too far from each other which will cause a huge problem when deciding to put these dispensers. Since some neighbourhoods which has more venues will overuse the dispensers as others will not be used too much.



Conclusion:

- This project would be far better if we could get the density of the population in each neighborhood and so we might have gotten better clustering and a more realistic one. However, i wasn't able to find data which provides the population of each of the neighborhoods in our lab dataset, so instead i tried to think of a different way to do so.
- Its that the KNN clustering fails to perform well on certain neighborhoods. However, the results obtained are still satisfying.